



# Integrated molecular landscape of Parkinson's disease

## Citation

Klemann, C. J. H. M., G. J. M. Martens, M. Sharma, M. B. Martens, O. Isacson, T. Gasser, J. E. Visser, and G. Poelmans. 2017. "Integrated molecular landscape of Parkinson's disease." NPJ Parkinson's Disease 3 (1): 14. doi:10.1038/s41531-017-0015-3. <http://dx.doi.org/10.1038/s41531-017-0015-3>.

## Published Version

doi:10.1038/s41531-017-0015-3

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33490802>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

## ARTICLE OPEN

## Integrated molecular landscape of Parkinson's disease

C. J. H. M. Klemann<sup>1</sup>, G. J. M. Martens<sup>1</sup>, M. Sharma<sup>2</sup>, M. B. Martens<sup>3</sup>, O. Isacson<sup>4</sup>, T. Gasser<sup>5</sup>, J. E. Visser<sup>1,6,7</sup> and G. Poelmans<sup>1,8,9</sup>

Parkinson's disease is caused by a complex interplay of genetic and environmental factors. Although a number of independent molecular pathways and processes have been associated with familial Parkinson's disease, a common mechanism underlying especially sporadic Parkinson's disease is still largely unknown. In order to gain further insight into the etiology of Parkinson's disease, we here conducted genetic network and literature analyses to integrate the top-ranked findings from thirteen published genome-wide association studies of Parkinson's disease (involving 13,094 cases and 47,148 controls) and other genes implicated in (familial) Parkinson's disease, into a molecular interaction landscape. The molecular Parkinson's disease landscape harbors four main biological processes—oxidative stress response, endosomal-lysosomal functioning, endoplasmic reticulum stress response, and immune response activation—that interact with each other and regulate dopaminergic neuron function and death, the pathological hallmark of Parkinson's disease. Interestingly, lipids and lipoproteins are functionally involved in and influenced by all these processes, and affect dopaminergic neuron-specific signaling cascades. Furthermore, we validate the Parkinson's disease–lipid relationship by genome-wide association studies data-based polygenic risk score analyses that indicate a shared genetic risk between lipid/lipoprotein traits and Parkinson's disease. Taken together, our findings provide novel insights into the molecular pathways underlying the etiology of (sporadic) Parkinson's disease and highlight a key role for lipids and lipoproteins in Parkinson's disease pathogenesis, providing important clues for the development of disease-modifying treatments of Parkinson's disease.

npj Parkinson's Disease (2017)3:14; doi:10.1038/s41531-017-0015-3

## INTRODUCTION

Parkinson's disease (PD) is the second most common neurodegenerative disease, with an estimated prevalence of 0.3%, affecting 1–2% of people over 60 years of age.<sup>1,2</sup> The pathological hallmark of PD is loss of dopaminergic (DA) neurons in the substantia nigra (SN), and the presence of protein aggregates (i.e., Lewy bodies) involving synuclein alpha (SNCA) in the residual DA neurons.<sup>3</sup> A number of biological processes that contribute to the pathogenesis of PD have been identified, including defects in mitochondrial function,<sup>4</sup> oxidative stress,<sup>5</sup> and protein aggregation.<sup>6–8</sup> However, detailed insights into the molecular mechanisms underlying these processes, and how they interact with each other, are essentially lacking. In many studies exploring PD pathogenesis, familial PD genes served as starting point. Thus far, at least eighteen genetic loci for familial PD have been found, and twelve familial PD candidate genes have been identified (*ATP13A2*, *DJ-1*, *DNAJC6*, *EIF4G1*, *FBXO7*, *LRRK2*, *PARK2*, *PINK1*, *PLA2G6*, *SNCA*, *SYNJ1*, and *VP535*) (refs 9, 10). However, as a mutation in one of these familial genes is found in only 5–10% of the cases, PD should be considered a predominantly sporadic disease,<sup>11,12</sup> with both genetic and environmental contributing risk factors. In recent years, 15 genome-wide association studies (GWASs) have investigated genetic risk factors for sporadic PD<sup>13–26</sup> but the functional coupling of the proteins encoded by the GWAS-identified

candidate genes to PD pathophysiology is often not clear. In the present study, we aimed to identify the core mechanisms underlying PD pathogenesis by using bioinformatics and extensive literature analyses to integrate (1) the genes corresponding to the top-ranked single-nucleotide polymorphisms (SNPs) found in published GWASs of sporadic PD, and (2) other PD candidate genes (e.g., familial PD genes) into a protein interaction landscape. This molecular landscape allowed us to identify the specific biological processes that are key in PD pathogenesis and provides clues for the development of novel PD treatment strategies.

## RESULTS

Selected PD GWAS genes and genetic network enrichment analysis

Thirteen of the fifteen published PD GWASs met our inclusion criteria (Supplementary Table 1) and were used to select a total of 451 PD GWAS candidate genes based on SNPs with  $p < 0.0001$  (Supplementary Table 2). Of the five most significantly enriched ingenuity pathway analysis (IPA) networks (Supplementary Table 3), the network with the highest enrichment score ( $p = 1.00E-44$ ) and the highest number of PD GWAS candidate gene-encoded proteins (28 proteins) served as the starting point for the building of the molecular landscape (Supplementary Fig. 1).

<sup>1</sup>Department of Molecular Animal Physiology, Donders Institute for Brain, Cognition and Behaviour, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University, Nijmegen, The Netherlands; <sup>2</sup>Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen, Tübingen, Germany; <sup>3</sup>Department of Neuroinformatics, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands; <sup>4</sup>Neuroregeneration Research Institute, McLean Hospital/Harvard Medical School, Belmont, MA, USA; <sup>5</sup>Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research, University of Tübingen, and German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany; <sup>6</sup>Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands; <sup>7</sup>Department of Neurology, Amphia Hospital, Breda, The Netherlands; <sup>8</sup>Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands and <sup>9</sup>Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence: G. Poelmans (Geert.poelmans@radboudumc.nl)

J. E. Visser and G. Poelmans contributed equally to this work.

Received: 11 August 2016 Revised: 5 February 2017 Accepted: 28 February 2017

Published online: 10 April 2017

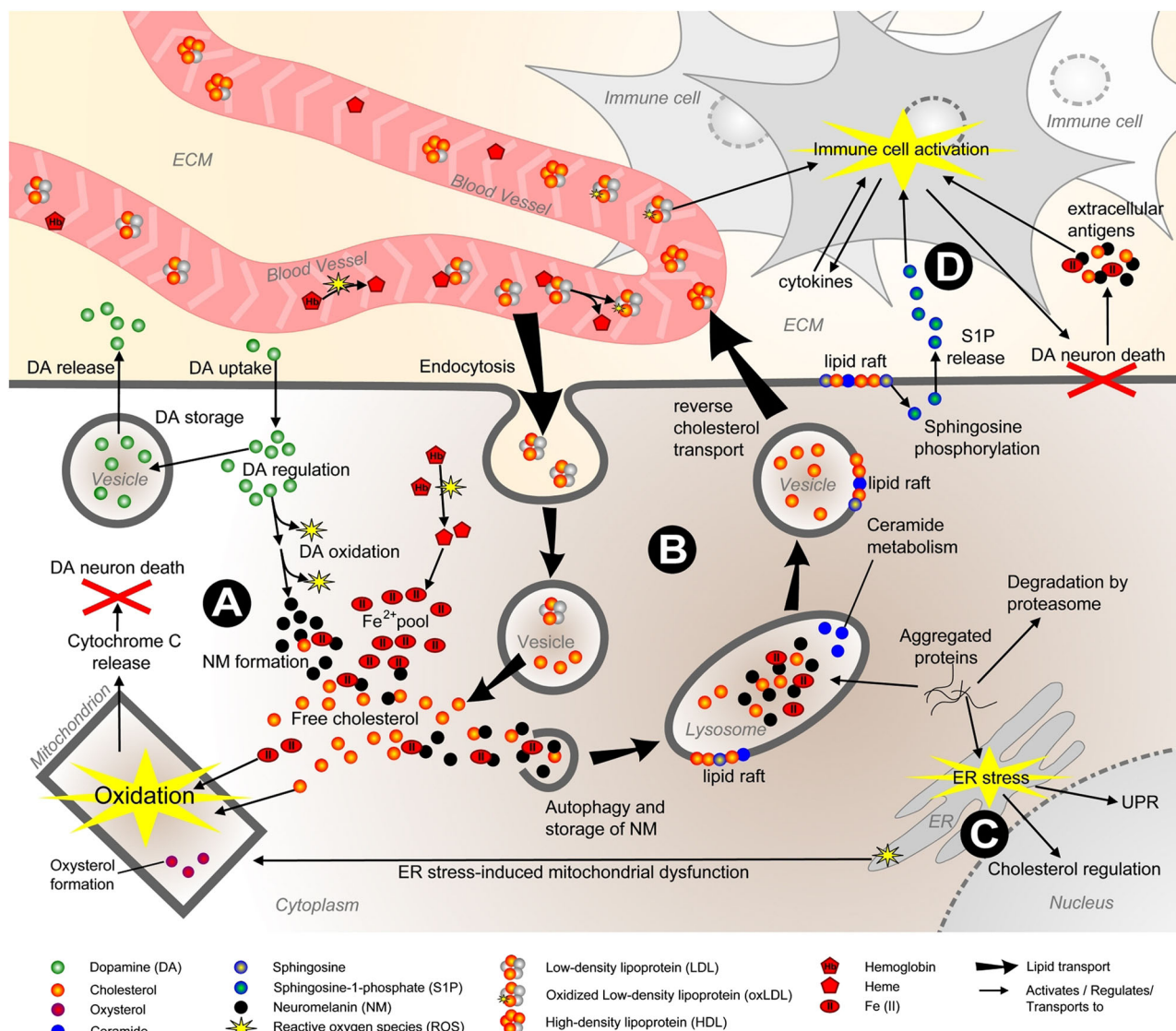
## The molecular landscape of PD

Guided by the most significantly enriched genetic network and extensive literature searches, we built a molecular landscape consisting of 260 interacting proteins (i.e., encoded by approximately 58% of the 451 PD GWAS genes, Supplementary Table 2), 128 proteins implicated in PD through other evidence (Supplementary Table 4) and 49 proteins that have not been directly linked to PD (yet) but have multiple functional interactions within the landscape (Supplementary Table 4). Approximately one in three landscape proteins are implicated in PD etiology through at least two types of evidence.

Supplementary Figures 2 and 3 show all relevant protein interactions in the PD landscape that are functionally involved in four main biological processes: oxidative stress response, endosomal-lysosomal functioning, endoplasmic reticulum (ER) stress response, and neuron death and immune response. The Supplementary Information provides a detailed and referenced description of the evidence linking all the proteins in the landscape. In Supplementary Table 5, we have indicated in which process(es) each landscape protein exerts its main effect and where it is located in Supplementary Figs 2 and/or 3.

The above being said, we here give a succinct description of the four main biological processes and signaling cascades in the PD landscape that are depicted in Fig. 1. Central in the landscape is signaling involving lipoproteins—i.e., low-density lipoprotein (LDL), high-density lipoprotein (HDL) and very low-density lipoprotein (VLDL)—and their component lipids and metabolites (e.g., cholesterol, oxysterols, sphingolipids such as ceramide and sphingosine, and triglycerides). Lipid and lipoprotein signaling represents the “common denominator” that functionally integrates, regulates and is regulated by the four landscape processes (Fig. 1a–d). Either by themselves or in combination, deficits or impairments in any of these four processes—each composed of multiple signaling cascades—can contribute to the degeneration and ultimately death of DA neurons.

First, deficits or impairments in dopamine synthesis and—linked to this—iron metabolism can cause an increased oxidative stress response (Fig. 1a). Dopamine can be either taken up through active transport or is newly synthesized in neurons and can subsequently be re-released (through vesicular exocytosis), degraded or (auto-)oxidized into neuromelanin (NM). Further, like



**Fig. 1** Overview of the molecular landscape of PD. The four main biological processes in the PD landscape—oxidative stress response (Fig. 1a), endosomal-lysosomal functioning (Fig. 1b), endoplasmic reticulum (ER) stress response (Fig. 1c), and neuron death and immune response (Fig. 1d)—are depicted. ECM extracellular matrix, ER endoplasmic reticulum, UPR unfolded protein response

erythrocytes (see below), SN DA neurons have a high-oxygen demand and express oxygen-carrying hemoglobin.

Through oxidation, cytotoxic heme is released from hemoglobin and then converted in DA neurons to ferrous iron, Fe(II). Fe(II) increases oxidative stress and together with free cholesterol—that is taken up by neurons through lipoproteins (see below)—induces mitochondrial oxysterol formation. In turn, this causes mitochondrial dysfunction and triggers the release of pro-apoptotic cytochrome c and, eventually, neuron death.

The second main landscape process centers around the (dys) regulation of endosomal-lysosomal functioning (Fig. 1b). Neuronal uptake of cholesterol occurs through the endosomal system, i.e., after neuronal uptake through vesicular endocytosis, LDL particles are processed into their composite parts: proteins, free cholesterol and other lipids. Free cholesterol and Fe(II) are bound in complexes by NM, which are then stored in lysosomes through autophagy. Hence, NM complex formation prevents the above described Fe(II)- and cholesterol-induced oxidative stress response. Moreover, their ageing-related increase in NM content and the associated increased demands on lysosomal function renders DA neurons particularly vulnerable to lysosomal defects. Other important lysosomal functions include the degradation of misfolded or aggregated proteins (such as pathological SNCA aggregates), the regulation of ceramide metabolism and reverse cholesterol transport, i.e., the vesicle-mediated transport and exocytosis of cholesterol into HDL particles in the bloodstream (and back to the liver). As such, a defect in any of these endosomal-lysosomal system components results in disturbed levels of lipids such as cholesterol and ceramide. In turn, these disturbed lipid levels affect membrane function in general and more specifically the functioning of so-called lipid rafts—microdomains of the vesicular, lysosomal, and plasma membrane containing high amounts of cholesterol and sphingolipids and crucial for membrane function—and hence processes such as autophagy, endo- and exocytosis. Deficient lysosomal function together with reduced degradation by the proteasome also leads to misfolded or aggregated protein formation.

Misfolded/aggregated proteins trigger the ER stress response (Fig. 1c), the third main landscape process, and subsequent activation of the protective unfolded protein response (UPR) as well as stimulation of cholesterol influx through upregulating the expression of key lipoprotein receptors. Prolonged ER stress that can no longer be counteracted by the UPR induces mitochondrial dysfunction, which eventually results in DA neuron death.

Lastly, apart from or in addition to dysregulated processes within DA neurons (as described above), DA neuron death can be the consequence of external factors such as an exaggerated immune response (Fig. 1d), the fourth landscape process. In this respect, immune cells are activated and attracted to damaged or already dying DA neurons by extracellular factors such as the sphingolipid-derived sphingosine-1-phosphate (S1P), triglyceride-rich extraneuronal VLDL particles (not shown), heme-oxidized LDL (oxLDL, see above), and various cytokines. Subsequently, the damaged/dying DA neurons are removed by the activated immune cells, an essentially normal and adequate response that is exaggerated in PD by DA neuron-specific antigens such as SNCA aggregates and NM complexes—released by dying DA neurons—creating a vicious cycle of DA neuron death and immune cell activation.

#### Polygenic risk score (PRS) analyses

Because our molecular landscape pointed towards an important role for lipids and lipoproteins in PD etiology (see above), we conducted PRS analyses using the tool PRSice,<sup>27</sup> with GWAS data for the blood levels of various lipids and lipoproteins<sup>28</sup> as base samples and meta-analytic PD GWAS data from the International Parkinson Disease Genomics Consortium (IPDGC)<sup>29</sup> as target

sample. We found statistically significant evidence (false discovery rate (FDR)-corrected  $p < 0.05$ ) for a shared genetic etiology between the lipid traits “total cholesterol levels” and “total triglyceride levels” and PD, with the most predictive  $p$ -value threshold ( $p_T$ ) at 0.001 and 0.05, respectively (Fig. 2). In contrast, the lipoprotein traits “total HDL levels” and “total LDL levels” yielded no evidence for a shared genetic risk with PD (Fig. 2). For the various combinations of increased or decreased HDL and LDL levels, we found significant evidence for a shared genetic etiology between PD and the combined trait “increased HDL + increased LDL” (most predictive  $p_T = 0.05$ ) (Supplementary Fig. 4).

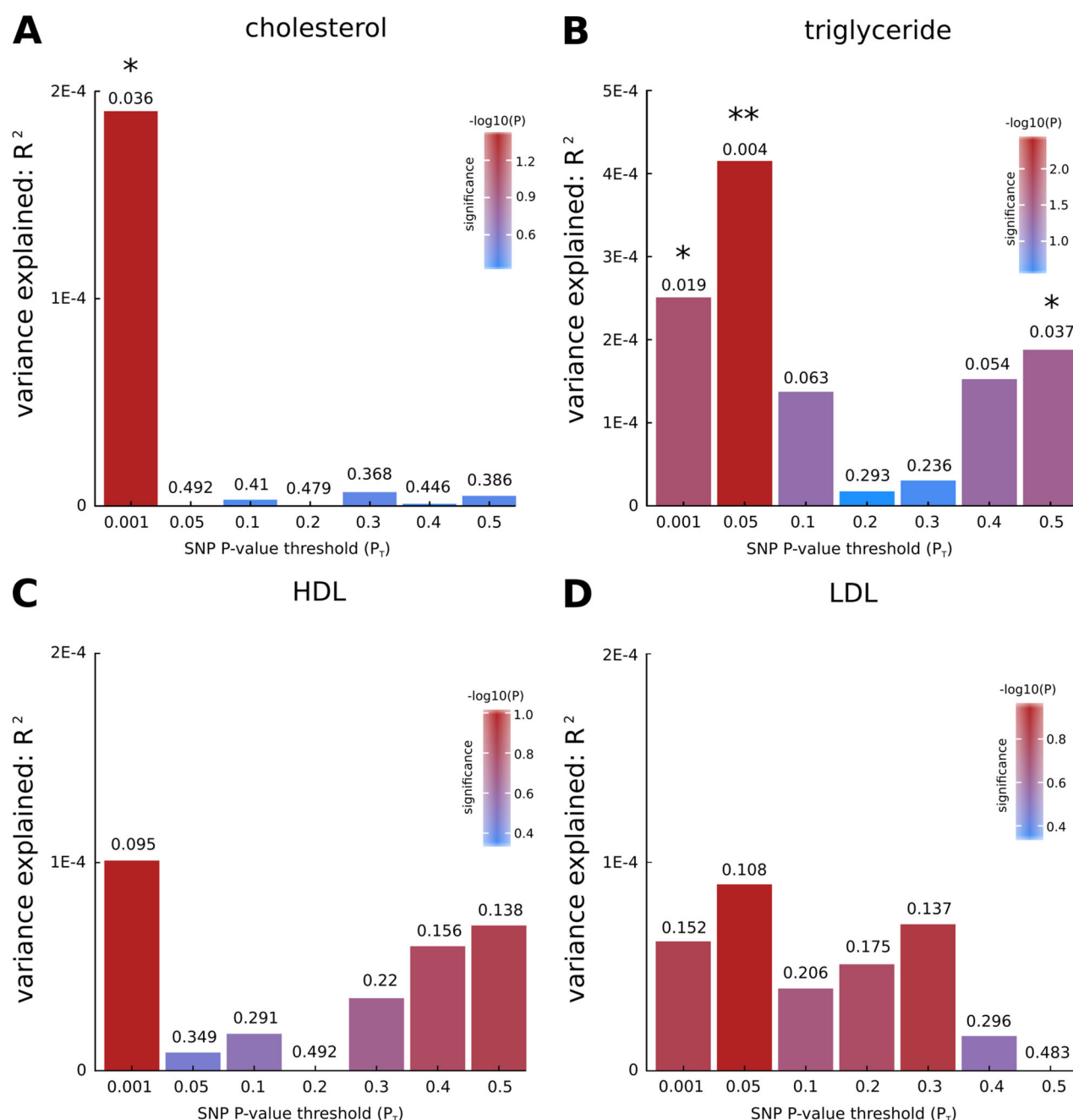
#### DISCUSSION

In this study, we integrated the available genetic and expression data with data about environmental risk factors into a protein interaction landscape that reveals the main biological processes and signaling cascades that are affected in PD and occur in DA neurons and immune cells. Our PD landscape represents processes and cascades that are affected in both the monogenic, familial and the more prevalent polygenic, sporadic forms of PD. In this respect, the molecular landscape includes the “classic” processes and cascades known to be affected in PD that are based on the familial PD candidate genes (e.g., SNCA, PARK2, LRRK2): mitochondrial function, oxidative stress and protein aggregation. In addition, the landscape harbors more novel processes that have been less well studied in relation to PD pathogenesis yet, such as impairments in lysosomal function and immune response regulation. The landscape does not imply a “sequence of events” that leads to DA neuron loss, i.e., a number of (impaired) biological processes and cascades that occur in a temporally and/or spatially distinct order. Instead, deficits in any of the main landscape processes/cascades, either by themselves or in combination with deficits of other processes/cascades, may cause DA neurons to die. Moreover, an aging-related decline in the functionality and/or efficiency of landscape processes/cascades may play a role in PD onset and progression. For example, a gradual buildup of NM or aggregated proteins may disturb lysosomal function in DA neurons<sup>30</sup> or an age-related decrease in the expression and activity of ER folding enzymes can compromise proper protein folding.<sup>31</sup>

Lipid and lipoprotein signaling functionally integrates, regulates and is regulated by the key landscape processes and cascades. Any disturbance of these processes and cascades can (eventually) result in DA neuron death, which is further aggravated or initiated by an increased (auto-)immune response.<sup>32</sup> The involvement of deficient lipid and lipoprotein signaling in PD pathophysiology is corroborated by a number of environmental risk factor studies. Increased plasma levels of total cholesterol are associated with a lower PD risk.<sup>33–35</sup> Nevertheless, a recent meta-analysis did not show an effect of higher or lower dietary cholesterol intake on PD risk,<sup>36</sup> suggesting that direct cholesterol intake through food may not play a major role in PD etiology. Further, low plasma levels of LDL are linked to a higher PD risk,<sup>37, 38</sup> whereas high-plasma HDL and CSF oxysterol levels are associated with increased PD risk and duration.<sup>39, 40</sup> In addition, the levels of oxidized LDL, oxysterols and sphingolipids are increased in the plasma of PD patients.<sup>41–43</sup> Thus, PD patients have a lower LDL:HDL ratio that is associated with a lower risk of cardiovascular disease (CVD)<sup>44</sup> and could at least to some extent explain why the PD population is indeed less susceptible to developing CVD.<sup>39</sup> Apart from the observed dysregulated levels of cholesterol (metabolites) and cholesterol-containing lipoproteins in PD patients, lower serum levels of triglycerides—which are highly enriched in VLDL particles—associate with an increased PD risk.<sup>45, 46</sup>

Intriguingly, we found a significant overlap between the polygenic risk associated with total cholesterol and triglyceride levels and PD. We also identified a shared genetic etiology





**Fig. 2** Bar plots from PRSice for shared genetic etiology between four lipid and lipoprotein traits (cholesterol, triglyceride, HDL, and LDL levels) and PD showing results at broad  $p$ -value thresholds ( $P_T$ ). The numbers above the bars indicate the  $p$ -values for shared genetic etiology, and these  $p$ -values were corrected using the false discovery rate (FDR) method; \*denotes FDR-corrected  $p < 0.05$ , \*\*denotes FDR-corrected  $p < 0.01$

between the combined lipoprotein trait “increased HDL + increased LDL” and PD. To our knowledge, we are the first to find a shared genetic risk between quantitative traits and a neurodegenerative disease but these findings need to be replicated in larger data sets, especially for the target sample, i.e., the PD GWAS data set. Together, the epidemiological and our PRS analysis findings indicate that the link between specific lipid/lipoprotein traits and PD may be the result of both shared environmental and genetic risk factors.

Given the converging evidence for lipid and lipoprotein signaling playing a key role in PD etiology, compounds that modulate lipid/lipoprotein levels could represent effective novel PD treatments. In this respect, statins—inhibitors of peripheral

cholesterol synthesis that are used to treat hypercholesterolemia and hence prevent cholesterol-associated CVD—have a neuro-protective effect in the rat brain,<sup>47</sup> but their effect on PD risk remains unclear.<sup>48–55</sup> Interestingly, the only published prospective study that has adjusted for baseline cholesterol levels before statin treatment has found that statin use is associated with a significantly higher PD risk,<sup>35</sup> which is in keeping with the observation that higher total plasma cholesterol levels—which are lowered by statins—are protective against PD. Other signaling molecules from the landscape that affect cholesterol and lipoprotein levels are testosterone and vitamin D3. Caucasian male PD patients show significantly reduced testosterone levels,<sup>56–58</sup> and free testosterone levels are positively correlated

with LDL, HDL, and total cholesterol levels.<sup>59</sup> Therefore, decreased testosterone levels may impact on several key PD landscape processes, as testosterone regulates the efflux of LDL and HDL to the circulation.<sup>60</sup> Hence, testosterone could be used for treating PD in male patients and indeed, testosterone treatment has some modest beneficial effects in men with PD.<sup>61, 62</sup> Deficiency of vitamin D3—which affects cholesterol metabolism through down-regulating SREBF1 (ref. 63), the main transcriptional activator of lipid homeostasis and key landscape protein—has been consistently associated with an increased PD risk<sup>64</sup> and its supplementation may stabilize PD symptoms.<sup>65</sup>

Lastly, a number of landscape proteins that both regulate lipid/lipoprotein signaling and landscape cascades involved in DA neuron death represent attractive (novel) drug targets for PD. Examples include HMOX1 that prevents oxidative stress by heme, PSAP and its receptor GPR37 that mediate ceramide metabolism, the immunity-related ICAM1 that is regulated by extracellular lipids and (oxidized) lipoproteins, and plasmin that regulates the degradation of extracellular SNCA and (lipo)proteins.

In conclusion, our integrated molecular landscape yields detailed insights into the mechanisms underlying PD pathogenesis, and highlights the involvement of deficient lipid and lipoprotein signaling. These findings warrant future rigorous perturbation experiments in PD cell and animal models that may eventually provide validated drug target “leads” for the development of novel disease-modifying PD treatments.

## METHODS

### PD GWAS gene selection

The first step of our molecular landscape building approach<sup>66–68</sup> is the selection of candidate genes based on GWAS SNPs and their corresponding *p*-values. All 15 PD GWASs published to date were considered. Criteria for inclusion were a publicly available GWAS discovery sample with all SNPs associated at *p* < 0.0001. From the GWASs for which these data were available, we then selected the SNPs that were associated with PD at *p* < 0.0001 to compile a list of associated genes. The selected genes either contained a SNP that was located within an exonic, intronic or untranslated region of the gene or were found within 100 kilobases (kb) downstream and upstream of the SNP. The latter was based on the fact that the vast majority of expression quantitative trait loci (eQTL) for a given gene are located within 100 kb downstream and/or upstream of a gene<sup>69–71</sup> and because trait-associated SNPs are more likely to be eQTL.<sup>72</sup> The chosen cutoff for association (*p* < 0.0001) is often employed to designate “suggestive” association and has been used in GWASs of multiple disorders.<sup>73–75</sup> Subsequently, the literature was searched for additional (genetic) evidence linking the selected GWAS candidate genes to PD.

### Genetic network enrichment analysis

To identify enriched protein networks in the PD GWAS candidate genes, a network analysis using the IPA software package (<http://www.ingenuity.com>) was performed with default parameters, i.e., the analysis used the so-called reference set of known genes and endogenous chemicals. This reference set is accessible through the “Ingenuity Knowledge Base”, a repository of data from publicly accessible databases (e.g., BioGRID, IntAct) and data that are manually curated by Ingenuity through systematically reviewing published literature. In this respect, of all protein–protein interactions between the PD candidate proteins in the most significantly enriched IPA network shown in Supplementary Fig. 1 (see Results), we found approximately 4/5 through publicly accessible databases and 1/5 by IPA. In addition, only functional relationships that are corroborated by experimental evidence were included in the IPA networks. For each network, the Ingenuity software also generates an enrichment score that takes into account the number of eligible molecules/proteins in the network and its size, as well as the total number of network-eligible molecules analyzed and the total number of molecules in the Ingenuity Knowledge Base that could potentially be included in networks. This score is the negative logarithm of the right-tailed Fisher’s exact test result.

## Molecular landscape building

Following the network enrichment analysis, the literature was extensively searched for the functions and interactions of all proteins encoded by the candidate genes implicated through PD GWASs as well as other PD candidates implicated via other evidence, including genetic association studies, messenger RNA/protein expression studies and/or functional studies. First, we used the UniProt Protein Knowledge Base (<http://www.uniprot.org/uniprot>)<sup>76</sup> to gather basic information on the functions of all candidate genes and their encoded proteins. Subsequently and starting with the interactions in the most enriched genetic network, we used PubMed (<http://www.ncbi.nlm.nih.gov/sites/entrez>) to search for all functional, experimental evidence-based interactions between all PD candidate genes/proteins. While building the landscape in this way, we also included genes/proteins and metabolites that have no known link with PD, but have multiple—i.e., at least two different—functional interactions with PD-implicated proteins. Based on all gathered information, we generated a protein interaction landscape. The figures depicting this landscape were made using the drawing program Serif DrawPlus version 4.0 ([www.serif.com](http://www.serif.com)).

## PRS analyses

Our molecular landscape pointed towards an important role for lipids and lipoproteins in PD etiology (see below). Therefore, we conducted PRS analyses using the tool PRSice,<sup>27</sup> with summary statistics data from GWASs of blood levels of total cholesterol, total triglycerides, total HDL and total LDL as “base” samples (GWAS data for 188577 European-ancestry individuals from the general population)<sup>28</sup> and summary statistics data from a meta-analytic PD GWAS by the IPDGC as “target” sample (GWAS data for 5333 PD cases and 12,019 healthy control subjects, all of European ancestry)<sup>29</sup>. Using the default settings in PRSice we calculated the shared genetic etiology between the four lipid/lipoprotein traits and PD at seven broad *p*-value thresholds (indicated by *p<sub>T</sub>*), which were used to select the SNPs from the base sample that were included in the PRS analysis, i.e., *p<sub>T</sub>* < 0.001, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5. As such, the seven *p<sub>T</sub>* thresholds led to the selection of all SNPs that were associated with the base lipid/lipoprotein phenotype at *p* < 0.001, 0.05, 0.1 etc.

The calculated *p*-values indicating the significance of a shared genetic etiology between each lipid/lipoprotein trait and PD were aggregated and corrected for multiple comparisons using the FDR method, incorporating potential dependencies between *p*-values.<sup>77</sup>

To calculate the FDR, we used the mafdr function in MATLAB (R2012a, The Mathworks, Natick, MA, USA) using the bootstrap selection method for the FDR parameter lambda. FDR was set to not lower *p*-values below uncorrected *p*-values, which would have occurred due to overall (relatively) low uncorrected *p*-values.

In addition to calculating the shared genetic etiology between the four lipid/lipoprotein traits and PD, we performed similar analyses using four “combined lipoprotein traits”. In order to do this, we first divided the HDL and LDL level summary statistics GWAS data into four groups of SNPs, i.e., all SNPs associated with (1) increased HDL levels, (2) increased LDL levels, (3) decreased HDL levels, and (4) decreased LDL levels (where “increased” and “decreased” refer to all SNPs that had an effect size-indicating beta > 0 and beta < 0, respectively). Subsequently, we conducted PRSice analyses with four combined data sets as base sample, i.e., (1) all SNPs increasing HDL and LDL levels, (2) all SNPs decreasing HDL and LDL levels, (3) all SNPs increasing HDL and decreasing LDL levels, and (4), all SNPs decreasing HDL and increasing LDL levels. Before the data sets were fed into PRSice, an equivalent, weighted *p*-value for each SNP in each of the four combined HDL/LDL data sets was calculated as:  $\frac{2}{p_{eq}} = \frac{1}{p_{HDL}} + \frac{1}{p_{LDL}}$

## ACKNOWLEDGEMENTS

We are particularly grateful to the International Parkinson Disease Genomics Consortium (IPDGC)—of which authors T. Gasser and M. Sharma are members—for providing us with the summary statistics GWAS data from the discovery sample of their meta-analytic GWAS of PD<sup>29</sup> that we used as the target sample for our PRS analyses. J.E. Visser was supported by Stichting Parkinsonfonds, the Netherlands Organization for Scientific Research (NWO/ZonMw, VENI 916.12.167) and The Netherlands Brain Foundation (F2014(1)-16). T. Gasser and M. Sharma received funding through the EU Joint Programme—Neurodegenerative Disease Research (JPND) project “Courage-PD”. This project is supported through the following funding organizations under the aegis of JPND ([www.jpnd.eu](http://www.jpnd.eu)): France, the French National Research Agency; Germany, the German Bundesministerium für Bildung und Forschung; Israel, the Israeli Ministry of Health; Italy, the Italian Ministry of Health/

Ministry of Education, Universities and Research; Luxembourg, the Luxembourgian National Research Fund; The Netherlands, the Netherlands Organization for Health Research and Development; Norway, the Research Council of Norway; Portugal, the Portuguese Foundation for Science and Technology; Spain, the Spanish National Institute of Health Carlos III; United Kingdom, the Medical Research Council. O. Isacson received funding from the National Institutes of Health grant R21NS084149, the Michael J Fox Foundation and the Consolidated Anti-Aging Foundation.

## AUTHOR CONTRIBUTIONS

C.K. integrated the available data into the PD molecular landscape. C.K., G.M., J.V., and G.P. were involved in designing the study. C.K., G.M., O.J., T.G., J.V., G.P. have contributed to the interpretation of the results. M.S. and T.G. have provided the summary statistics GWAS data for the polygenic risk score analysis that was performed by M.M. All authors read and approved the final manuscript.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## REFERENCES

- Nussbaum, R. L. & Ellis, C. E. Alzheimer's disease and Parkinson's disease. *N. Engl. J. Med.* **348**, 1356–1364 (2003).
- de Lau, L. M. & Breteler, M. M. Epidemiology of Parkinson's disease. *Lancet Neurol.* **5**, 525–535 (2006).
- Gibb, W. R. & Lees, A. J. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry.* **51**, 745–752 (1988).
- Ryan, B. J., Hoek, S., Fon, E. A. & Wade-Martins, R. Mitochondrial dysfunction and mitophagy in Parkinson's: from familial to sporadic disease. *Trends Biochem. Sci.* **40**, 200–210 (2015).
- Blesa, J., Trigo-Damas, I., Quiroga-Varela, A. & Jackson-Lewis, V. R. Oxidative stress and Parkinson's disease. *Front. Neuroanat.* **9**, 91 (2015).
- Trojanowski, J. Q., Goedert, M., Iwatsubo, T. & Lee, V. M. Fatal attractions: abnormal protein aggregation and neuron death in Parkinson's disease and Lewy body dementia. *Cell Death. Differ.* **5**, 832–837 (1998).
- Gundersen, V. Protein aggregation in Parkinson's disease. *Acta Neurol. Scand. Suppl.* **190**, 82–87 (2010).
- Ebrahimi-Fakhari, D., Wahlster, L. & McLean, P. J. Protein degradation pathways in Parkinson's disease: curse or blessing. *Acta Neuropathol.* **124**, 153–172 (2012).
- Bonifati, V. Genetics of Parkinson's disease—state of the art, 2013. *Parkinsonism. Relat. Disord.* **20**, S23–S28 (2014).
- Spatola, M. & Wider, C. Genetics of Parkinson's disease: the yield. *Parkinsonism. Relat. Disord.* **20**, S35–S38 (2014).
- Thomas, B. & Beal, M. F. Parkinson's disease. *Hum. Mol. Genet.* **2**, R183–R194 (2007).
- Klein, C. & Westenberger, A. Genetics of Parkinson's disease. *Cold Spring Harb. Perspect. Med.* **2**, a008888 (2012).
- Maraganore, D. M. et al. High-resolution whole-genome association study of Parkinson disease. *Am. J. Hum. Genet.* **77**, 685–693 (2005).
- Fung, H. C. et al. Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol.* **5**, 911–916 (2006).
- Pankratz, N. et al. Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *Hum. Genet.* **124**, 593–605 (2009).
- Latourelle, J. C. et al. Genomewide association study for onset age in Parkinson disease. *BMC Med. Genet.* **10**, 98 (2009).
- Satake, W. et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* **41**, 1303–1307 (2009).
- Simon-Sanchez, J. et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **41**, 1308–1312 (2009).
- Edwards, T. L. et al. Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann. Hum. Genet.* **74**, 97–109 (2010).
- Hamza, T. H. et al. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat. Genet.* **42**, 781–785 (2010).
- Spencer, C. C. et al. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum. Mol. Genet.* **20**, 345–353 (2011).
- Saad, M. et al. Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population. *Hum. Mol. Genet.* **20**, 615–627 (2011).
- Simon-Sanchez, J. et al. Genome-wide association study confirms extant PD risk loci among the Dutch. *Eur. J. Hum. Genet.* **19**, 655–661 (2011).
- Do, C. B. et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* **7**, e1002141 (2011).
- Liu, X. et al. Genome-wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. *BMC Med. Genet.* **12**, 104 (2011).
- Hernandez, D. G. et al. Genome wide assessment of young onset Parkinson's disease from Finland. *PLoS ONE* **7**, e41859 (2012).
- Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: polygenic risk score software. *Bioinformatics* **31**, 1466–1468 (2015).
- Willer, C. J. et al. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013).
- Nalls, M. A. et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* **377**, 641–649 (2011).
- Isacson, O. Lysosomes to combat Parkinson's disease. *Nat. Neurosci.* **18**, 792–793 (2015).
- Salganik, M. et al. The loss of glucose-regulated protein 78 (GRP78) during normal aging or from siRNA knockdown augments human alpha-synuclein (alpha-syn) toxicity to rat nigral neurons. *Neurobiol. Aging* **36**, 2213–2223 (2015).
- Deleidi, M. & Isacson, O. Viral and inflammatory triggers of neurodegenerative diseases. *Sci. Transl. Med.* **4**, 121ps3 (2012).
- de Lau, L. M., Koudstaal, P. J., Hofman, A. & Breteler, M. M. Serum cholesterol levels and the risk of Parkinson's disease. *Am. J. Epidemiol.* **164**, 998–1002 (2006).
- Miyake, Y. et al. Case-control study of risk of Parkinson's disease in relation to hypertension, hypercholesterolemia, and diabetes in Japan. *J. Neurol. Sci.* **293**, 82–86 (2010).
- Huang, X. et al. Statins, plasma cholesterol, and risk of Parkinson's disease: a prospective study. *Mov. Disord.* **30**, 552–559 (2015).
- Wang, A., Lin, Y., Wu, Y. & Zhang, D. Macronutrients intake and risk of Parkinson's disease: a meta-analysis. *Geriatr. Gerontol. Int.* **15**, 606–616 (2015).
- Huang, X. et al. Lower low-density lipoprotein cholesterol levels are associated with Parkinson's disease. *Mov. Disord.* **22**, 377–381 (2007).
- Huang, X. et al. cholesterol and increased risk of Parkinson's disease: prospective results from Honolulu-Asia Aging Study. *Mov. Disord.* **23**, 1013–1018 (2008).
- Cassani, E. et al. Cardiometabolic factors and disease duration in patients with Parkinson's disease. *Nutrition* **29**, 1331–1335 (2013).
- Bjorkhem, I. et al. Oxysterols and Parkinson's disease: evidence that levels of 24S-hydroxycholesterol in cerebrospinal fluid correlates with the duration of the disease. *Neurosci. Lett.* **555**, 102–105 (2013).
- Andican, G. et al. Plasma oxidative and inflammatory markers in patients with idiopathic Parkinson's disease. *Acta Neurol. Belg.* **112**, 155–159 (2012).
- Seet, R. C. et al. Oxidative damage in Parkinson disease: Measurement using accurate biomarkers. *Free Radic. Biol. Med.* **48**, 560–566 (2010).
- Mielke, M. M. et al. Plasma ceramide and glucosylceramide metabolism is altered in sporadic Parkinson's disease and associated with cognitive impairment: a pilot study. *PLoS ONE* **8**, e73094 (2013).
- Jukema, J. W. et al. LDL-C/HDL-C ratio in subjects with cardiovascular disease and a low HDL-C: results of the RADAR (Rosuvastatin and Atorvastatin in different dosages and reverse cholesterol transport) study. *Curr. Med. Res. Opin.* **21**, 1865–1874 (2005).
- Wei, Q. et al. Reduced serum levels of triglyceride, very low density lipoprotein cholesterol and apolipoprotein B in Parkinson's disease patients. *PLoS ONE* **8**, e75743 (2013).
- Saaksjarvi, K., Knekt, P., Mannisto, S., Lyytinen, J. & Heliovaara, M. Prospective study on the components of metabolic syndrome and the incidence of Parkinson's disease. *Parkinsonism. Relat. Disord.* **21**, 1148–1155 (2015).
- He, X., Jenner, A. M., Ong, W. Y., Farooqui, A. A. & Patel, S. C. Lovastatin modulates increased cholesterol and oxysterol levels and has a neuroprotective effect on rat hippocampal neurons after kainate injury. *J. Neuropathol. Exp. Neurol.* **65**, 652–663 (2006).
- Wolozin, B. et al. Simvastatin is associated with a reduced incidence of dementia and Parkinson's disease. *BMC Med.* **5**, 20 (2007).
- Wahner, A. D., Bronstein, J. M., Bordelon, Y. M. & Ritz, B. Statin use and the risk of Parkinson disease. *Neurology* **70**, 1418–1422 (2008).
- Gao, X., Simon, K. C., Schwarzschild, M. A. & Ascherio, A. Prospective study of statin use and risk of Parkinson disease. *Arch. Neurol.* **69**, 380–384 (2012).
- Friedman, B., Lahad, A., Dresner, Y. & Vinker, S. Long-term statin use and the risk of Parkinson's disease. *Am. J. Manag. Care* **19**, 626–632 (2013).

52. Becker, C., Jick, S. S. & Meier, C. R. Use of statins and the risk of Parkinson's disease: a retrospective case-control study in the UK. *Drug Saf.* **31**, 399–407 (2008).
53. Samii, A., Carleton, B. C. & Etminan, M. Statin use and the risk of Parkinson disease: a nested case control study. *J. Clin. Neurosci.* **15**, 1272–1273 (2008).
54. Hippisley-Cox, J. & Coupland, C. Unintended effects of statins in men and women in England and Wales: population based cohort study using the QResearch database. *BMJ* **340**, c2197 (2010).
55. Ritz, B. et al. Statin use and Parkinson's disease in Denmark. *Mov. Disord.* **25**, 1210–1216 (2010).
56. Ready, R. E., Friedman, J., Grace, J. & Fernandez, H. Testosterone deficiency and apathy in Parkinson's disease: a pilot study. *J. Neurol. Neurosurg. Psychiatry* **75**, 1323–1326 (2004).
57. Kenangil, G., Orken, D. N., Ur, E., Forta, H. & Celik, M. The relation of testosterone levels with fatigue and apathy in Parkinson's disease. *Clin. Neurol. Neurosurg.* **111**, 412–414 (2009).
58. Brodacki, B. et al. cGMP level in idiopathic Parkinson's disease patients with and without cardiovascular disease - A pilot study. *Parkinsonism. Relat. Disord.* **17**, 689–692 (2011).
59. Chock, B., Lin, T. C., Li, C. S. & Swislocki, A. Plasma testosterone is associated with Framingham risk score. *Aging Male* **15**, 134–139 (2012).
60. Rubinow, K. B. et al. Acute sex steroid withdrawal increases cholesterol efflux capacity and HDL-associated clusterin in men. *Steroids* **77**, 454–460 (2012).
61. Okun, M. S. et al. Beneficial effects of testosterone replacement for the nonmotor symptoms of Parkinson disease. *Arch. Neurol.* **59**, 1750–1753 (2002).
62. Okun, M. S. et al. Testosterone therapy in men with Parkinson disease: results of the TEST-PD Study. *Arch. Neurol.* **63**, 729–735 (2006).
63. Kong, J. & Li, Y. C. Molecular mechanism of 1,25-dihydroxyvitamin D3 inhibition of adipogenesis in 3T3-L1 cells. *Am. J. Physiol. Endocrinol. Metab.* **290**, E916–E924 (2006).
64. Wang, L. et al. Vitamin D from different sources is inversely associated with Parkinson disease. *Mov. Disord.* **30**, 560–566 (2015).
65. Suzuki, M. et al. Randomized, double-blind, placebo-controlled trial of vitamin D supplementation in Parkinson disease. *Am. J. Clin. Nutr.* **97**, 1004–1013 (2013).
66. Poelmans, G., Pauls, D. L., Buitelaar, J. K. & Franke, B. Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. *Am. J. Psychiatry* **168**, 365–377 (2011).
67. Poelmans, G., Franke, B., Pauls, D. L., Glennon, J. C. & Buitelaar, J. K. AKAPs integrate genetic findings for autism spectrum disorders. *Transl. Psychiatry* **3**, e270 (2013).
68. van de Vondervoort, I. et al. An integrated molecular landscape implicates the regulation of dendritic spine formation through insulin-related signalling in obsessive-compulsive disorder. *J. Psychiatry Neurosci.* **41**, 140327 (2016).
69. Veyrieras, J. B. et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet.* **4**, e1000214 (2008).
70. Gherman, A., Wang, R. & Avramopoulos, D. Orientation, distance, regulation and function of neighbouring genes. *Hum. Genomics* **3**, 143–156 (2009).
71. Pickrell, J. K. et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* **464**, 768–772 (2010).
72. Nicolae, D. L. et al. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet.* **6**, e1000888 (2010).
73. Ma, D. et al. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann. Hum. Genet.* **73**, 263–273 (2009).
74. Xu, W. et al. Genome-wide association study of bipolar disorder in Canadian and UK populations corroborates disease loci including SYNE1 and CSMD1. *BMC Med. Genet.* **15**, 2 (2014).
75. Lindstrom, S. et al. Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nat. Commun.* **5**, 5303 (2014).
76. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, (Database issue): D204–D212 (2015).
77. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57.1**, 289–300 (1995).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the *npj Parkinson's Disease* website (doi:10.1038/s41531-017-0015-3).